



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Seed, Brian et al.
- (ii) TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES AND METHODS
- (iii) NUMBER OF SEQUENCES: 16
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Clark & Elbing LLP  
(B) STREET: 585 Commercial Street  
(C) CITY: Boston  
(D) STATE: MA  
(E) COUNTRY: USA  
(F) ZIP: 02109-1024
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: 08/756,018  
(B) FILING DATE: 25-NOV-96  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 08/661,960  
(B) FILING DATE: 12-JUN-1996
- (viii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: US 60/000,213  
(B) FILING DATE: 14-JUN-1995
- (ix) ATTORNEY/AGENT INFORMATION:  
(A) NAME: Elbing, Karen Lech  
(B) REGISTRATION NUMBER: 35,238  
(C) REFERENCE/DOCKET NUMBER: 00786/284002
- (x) TELECOMMUNICATION INFORMATION:  
(A) TELEPHONE: 617/723-6777  
(B) TELEFAX: 617/723-8962  
(C) TELEX:

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 amino acids  
(B) TYPE: amino acid

- (C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala Thr Glu Ala Gln Thr Thr Pro Pro Ala  
1                   5                   10

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Thr Asn Ser Leu Glu Thr Ser Thr Gly Thr Ser Gly Pro Pro  
1                   5                   10                   15

Val Thr

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gln Leu Trp Asp Thr Trp Ala Asp Glu Ala Glu Lys Ala Leu Gly Pro  
1                   5                   10                   15

Leu Leu Ala Arg Asp Arg Arg Gln Ala Thr Glu Tyr Glu Tyr Leu Asp  
20                   25                   30

Tyr Asp Phe Leu Pro Glu Thr Glu Pro Pro  
35                   40

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids

- (B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Arg Asp Arg Arg Gln Ala Thr Glu Tyr Glu Tyr Leu Asp Tyr Asp Phe  
1 5 10 15

Leu Pro Glu Thr  
20

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Arg Asp Arg Arg Gln Ala Thr Glu Phe Glu Phe Leu Asp Phe Asp Phe  
1 5 10 15

Leu Pro Glu Thr  
20

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Arg Asp Arg Arg Gln Ala Ala Glu Tyr Glu Tyr Leu Asp Tyr Asp Phe  
1 5 10 15

Leu Pro Glu Ala  
20

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: not relevant  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Arg Asp Arg Arg Gln Ala Ala Glu Phe Glu Phe Leu Asp Phe Asp Phe  
1 5 10 15

Leu Pro Glu Ala  
20

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2287 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAGCTTACCA CCATGGACTG GACCTGGAGG TTCC TCTTCT TTGTGGTGGC AGCAGCTACA 60  
GGTGTCCAGT CCCAGGTGCA GCTGGTGCAG TCTGGGGCTG AGGTGAAGAA GCCTGGGTCC 120  
TCGGTGAAGG TCTCCTGCAA GGCTTCTGGA GGCACCTTCA GCAGCTATGC TATCAGCTGG 180  
GTGCGACAGG CCCCTGGACA AGGGCTTGAG TGGATGGAG GGATCATCCC TATCTTTGGT 240  
ACAGCAAAC ACT CGCACAGAA GTTCCAGGGC AGAGTCACGA TTACCGCGGA CGAATCCACG 300  
AGCACAGCCT ACATGGAGCT GAGCAGCCTG AGATCTGAGG ACACGGCCGT GTATTACTGT 360  
GCGAGAGATA ATGGAGCGTA TTGTAGTGGT CGTAGCTGCT ACTCGGGCTG GTTCGACCCC 420  
TGGGGCCAGG GAACCCTGGT CACCGTCTCT TCAGGTGAGT ACTGAATTCT AGCTTTCTGG 480  
GGCAGGCCAG GCCTGACCTT GGCTTGGGG CAGGGAGGGG GCTAAGGTGA GGCAGGTGGC 540  
GCCAGCAGGT GCACACCCAA TGCCCATGAG CCCAGACACT GGACGCTGAA CCTCGCGGAC 600  
AGTTAAGAAC CCAGGGGCCT CTGCGCCTGG GCCCAGCTCT GTCCCACACC GCGGTCACAT 660  
GGCACCACT CTCTTGCAGC CTCCACCAAG GGCCCATCGG TCTTCCCCCT GGCACCCCTCC 720  
TCCAAGAGCA CCTCTGGGGG CACAGCGGCC CTGGGCTGCC TGGTCAAGGA CTACTTCCCC 780

GAACCGGTGA CGGTGTCGTG GAACTCAGGC GCCCTGACCA CGGGCGTGCA CACCTTCCCG 840  
 GCTGTCCTAC AGTCCTCAGG ACTCTACTCC CTCAGCAGCG TGGTGACCGT GCCCTCCAGC 900  
 AGCTTGGCA CCCAGACCTA CATCTGCAAC GTGAATCACA AGCCCAGCAA CACCAAGGTG 960  
 GACAAGAAAG TTGGTGAGAG GCCAGCACAG GGAGGGAGGG TGTCTGCTGG AAGCAGGCTC 1020  
 AGCGCTCCTG CCTGGACGCA TCCCGGCTAT GCACCCCCAG TCCAGGGCAG CAAGGCAGGC 1080  
 CCCGTCTGCC TCTTCACCCG GAGCCTCTGC CCGCCCCACT CATGCTCAGG GAGAGGGTCT 1140  
 TCTGGTTTT TCCCAGGCTC TGGGCAGGCA CAGGCTAGGT GCCCTTAACC CAGGCCCTGC 1200  
 ACACAAAGGG GCAGGTGCTG GGCTCAGACC TGCCAAGAGC CATATCCGGG AGGACCCCTGC 1260  
 CCCTGACCTA AGCCCACCCC AAAGGCCAAA CTCTCCACTC CCTCAGCTCG GACACCTTCT 1320  
 CTCCTCCAG ATTCCAGTAA CTCCCAATCT TCTCTCTGCA GAGCCCAAAT CTTGTGACAA 1380  
 AACTCACACA TGCCCACCGT GCCCAGGTAA GCCAGCCAG GCCTCGCCCT CCAGCTCAAG 1440  
 GCGGGACAGG TGCCCTAGAG TAGCCTGCAT CCAGGGACAG GCCCCAGCCG GGTGCTGACA 1500  
 CGTCCACCTC CATCTCTTCC TCAGCACCTG AACTCCTGGG GGGACCGTCA GTCTCCTCT 1560  
 TCCCCCCTAA ACCCAAGGAC ACCCTCATGA TCTCCGGAC CCTGAGGTC ACATGCGTGG 1620  
 TGGTGGACGT GAGCCACGAA GACCCTGAGG AAAGTTCAA CTGGTACGTG GACGGCGTGG 1680  
 AGGTGCATAA TGCCAAGACA AAGCCGCGGG AGGAGCAGTA CAACAGCAG TACCGGGTGG 1740  
 TCAGCGTCCT CACCGTCCTG CACCAGGACT GGCTGAATGG CAAGGAGTAC AAGTGAAGG 1800  
 TCTCCAACAA AGCCCTCCCA GCCCCATCG AGAAAACCCT CTCCAAAGCC AAAGGTGGGA 1860  
 CCCGTGGGGT GCGAGGGCCA CATGGACAGA GGCCGGCTCG GCCCACCCTC TGCCCTGAGA 1920  
 GTGACCGCTG TACCAACCTC TGTCTACAG GGCAGCCCCG AGAACACACAG GTGTACACCC 1980  
 TGCCCCCATC CCGGGATGAG CTGACCAAGA ACCAGGTCAAG CCTGACCTGC CTGGTCAAAG 2040  
 GCTTCTATCC CAGCGACATC GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACT 2100  
 ACAAGACCAAC GCCTCCCGTG CTGGACTCCG ACGGCTCCTT CTTCTCTAC AGCAAGCTCA 2160  
 CCGTGGACAA GAGCAGGTGG CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG 2220  
 CTCTGCACAA CCACTACACG CAGAAGAGCC TCTCCCTGTC TCCGGTAAA TGAGTGCAC 2280  
 GGCCGGC 2287

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 442 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Lys Leu Thr Thr Met Asp Trp Thr Trp Arg Phe Leu Phe Phe Val Val  
1 5 10 15

Ala Ala Ala Thr Gly Val Gln Ser Gln Val Gln Leu Val Gln Ser Gly  
20 25 30

Ala Glu Val Lys Lys Pro Gly Ser Ser Val Lys Val Ser Cys Lys Ala  
35 40 45

Ser Gly Gly Thr Phe Ser Ser Tyr Ala Ile Ser Trp Val Arg Gln Ala  
50 55 60

Pro Gly Gln Gly Leu Glu Trp Met Gly Gly Ile Ile Pro Ile Phe Gly  
65 70 75 80

Thr Ala Asn Tyr Ala Gln Lys Phe Gln Gly Arg Val Thr Ile Thr Ala  
85 90 95

Asp Glu Ser Thr Ala Arg Asp Asn Gly Ala Tyr Cys Ser Gly Gly Ser  
100 105 110

Cys Tyr Ser Gly Trp Phe Asp Pro Trp Gly Gln Gly Thr Leu Val Thr  
115 120 125

Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro  
130 135 140

Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val  
145 150 155 160

Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala  
165 170 175

Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly  
180 185 190

Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Asp Lys  
195 200 205

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys  
210 215 220

08756048 - 032067  
Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro  
225 230 235 240  
Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys  
245 250 255  
Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp  
260 265 270  
Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu  
275 280 285  
Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu  
290 295 300  
His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn  
305 310 315 320  
Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly  
325 330 335  
Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu  
340 345 350  
Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr  
355 360 365  
Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn  
370 375 380  
Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe  
385 390 395 400  
Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn  
405 410 415  
Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr  
420 425 430  
Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
435 440

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1894 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATGGCGCTGT CCTGGGTTCT TACAGTCCTG AGCCTCCTAC CTCTGCTGGA AGCCCAGATC 60  
 CCATTGTGTG CCAACCTAGT ACCGGTGCCC ATCACCAACG CCACCCTGGA CCAGATCACT 120  
 GGCAAGTGGT TTTATATCGC ATCGGCCTT CGAAACGAGG AGTACAATAA GTCGGTTTAG 180  
 GAGATCCAAG CAACCTTCTT TTACTTCACC CCCAACAAAGA CAGAGGACAC GATCTTCTC 240  
 AGAGAGTACC AGACCCGACA GGACCAGTGC ATCTATAACA CCACCTACCT GAATGTCCAG 300  
 CGGGAAAATG GGACCATCTC CAGATACGTG GGAGGCCAAG AGCATTGCG TCACTTGCTG 360  
 ATCCTCAGGG ACACCAAGAC CTACATGCTT GCTTTGACG TGAACGATGA GAAGAACTGG 420  
 GGGCTGTCTG TCTATGCTGA CAAGCCAGAG ACGACCAAGG AGCAACTGGG AGAGTTCTAC 480  
 GAAGCTCTCG ACTGCTTGCG CATTCCAAG TCAGATGTCG TGTACACCGA TTGGAAAAAG 540  
 GATAAGTGTG AGCCACTGGA GAAGCAGCAC GAGAAGGAGA GGAAACAGGA GGAGGGGGAA 600  
 TCGGATCCCG AGGGTGAGTA CTAAGCTTCA GCGCTCCTGC CTGGACGCAT CCCGGCTATG 660  
 CAGCCCCAGT CCAGGGCAGC AAGGCAGGCC CCGTCTGCCT CTTCACCCGG AGCCTCTGCC 720  
 CGCCCCACTC ATGCTCAGGG AGAGGGTCTT CTGGCTTTTT CCCAGGCTCT GGGCAGGCAC 780  
 AGGCTAGGTG CCCCTAACCC AGGCCCTGCA CACAAACGGG CAGGTGCTGG GCTCAGACCT 840  
 GCCAAGAGCC ATATCCGGGA GGACCCTGCC CCTGACCTAA GCCCACCCCA AAGGCCAAC 900  
 TCTCCACTCC CTCAGCTCGG ACACCTTCTC TCCTCCCAGA ~~TTCAGTAAC~~ TCCCAATCTT 960  
 CTCTCTGCAG AGCCCAAATC TTGTGACAAA ACTCACACAT GCCCACCGTG CCCAGGTAAG 1020  
 CCAGCCCAGG CCTCGCCCTC CAGCTCAAGG CGGGACAGGT GCCCTAGAGT ~~ACCTGCATC~~ 1080  
 CAGGGACAGG CCCCAGCCGG GTGCTGACAC GTCCACCTCC ATCTCTTCTT CAGCACCTGA 1140  
 ACTCCTGGGG GGACCGTCAG TCTTCCTCTT CCCCCCAAAA CCCAAGGACA CCCTCATGAT 1200  
 CTCCCGGACC CCTGAGGTCA CATGCGTGGT GGTGGACGTG AGCCACGAAG ACCCTGAGGT 1260  
 CAAGTTCAAC TGGTACGTGG ACGGCGTGG A ~~G~~ GGTGGCATAAT ~~G~~ GCCAAGACAA AGCCGCGGGA 1320  
 GGAGCAGTAC AACAGCACGT ACCGGGTGGT CAGCGTCCTC ACCGTCCTGC ACCAGGACTG 1380  
 GCTGAATGGC AAGGAGTACA AGTGCAAGGT CTCCAACAAA ~~G~~ CCCCTCCAG CCCCCATCGA 1440  
 GAAAACCATC TCCAAAGCCA AAGGTGGAC CCGTGGGGTG CGAGGGCCAC ATGGACAGAG 1500  
 GCCGGCTCGG CCCACCCCTCT GCCCTGAGAG TGACCGCTGT ACCAACCTCT GTCTACAGG 1560  
 GCAGCCCCGA GAACCACAGG TGTACACCCCT GCCCCCCATCC CGGGATGAGC TGACCAAGAA 1620

CCAGGTCAGC CTGACCTGCC TGGTCAAAGG CTTCTATCCC AGCGACATCG CCGTGGAGTG 1680  
GGAGAGCAAT GGGCAGCCGG AGAACAACTA CAAGACCACG CCTCCCGTGC TGGACTCCGA 1740  
CGGCTCCTTC TTCCTCTACA GCAAGCTCAC CGTGGACAAG AGCAGGTGGC AGCAGGGGAA 1800  
CGTCTTCTCA TGCTCCGTGA TGCATGAGGC TCTGCACAAC CACTACACGC AGAAGAGCCT 1860  
CTCCCTGTCT CCGGGTAAAT GAGTGCACG CCCG 1894

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 437 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Ala Leu Ser Trp Val Leu Thr Val Leu Ser Leu Leu Pro Leu Leu			
1	5	10	15
Glu Ala Gln Ile Pro Leu Cys Ala Asn Leu Val Pro Val Pro Ile Thr			
20	25	30	
Asn Ala Thr Leu Asp Gln Ile Thr Gly Lys Trp Phe Tyr Ile Ala Ser			
35	40	45	
Ala Phe Arg Asn Glu Glu Tyr Asn Lys Ser Val Gln Glu Ile Gln Ala			
50	55	60	

082256046 - 0220032

Thr Phe Phe Tyr Phe Thr Pro Asn Lys Thr Glu Asp Thr Ile Phe Leu  
65 70 75 80

Arg Glu Tyr Gln Thr Arg Gln Asp Gln Cys Ile Tyr Asn Thr Thr Tyr  
85 90 95

Leu Asn Val Gln Arg Glu Asn Gly Thr Ile Ser Arg Tyr Val Gly Gly  
100 105 110

Gln Glu His Phe Ala His Leu Leu Ile Leu Arg Asp Thr Lys Thr Tyr  
115 120 125

Met Leu Ala Phe Asp Val Asn Asp Glu Lys Asn Trp Gly Leu Ser Val  
130 135 140

Tyr Ala Asp Lys Pro Glu Thr Thr Lys Glu Gln Leu Gly Glu Phe Tyr  
145 150 155 160

Glu Ala Leu Asp Cys Leu Arg Ile Pro Lys Ser Asp Val Val Tyr Thr  
165 170 175

Asp Trp Lys Lys Asp Lys Cys Glu Pro Leu Glu Lys Gln His Glu Lys  
180 185 190

Glu Arg Lys Gln Glu Glu Gly Glu Ser Asp Pro Glu Gly Glu Pro Lys  
195 200 205

Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu  
210 215 220

Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr  
225 230 235 240

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val  
245 250 255

Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val  
260 265 270

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser  
275 280 285

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu  
290 295 300

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala  
305 310 315 320

Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro  
325 330 335

Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu Leu Thr Lys Asn Gln  
340 345 350

Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala  
355 360 365

Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr  
370 375 380

Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu  
385 390 395 400

Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser  
405 410 415

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser  
420 425 430

Leu Ser Pro Gly Lys  
435

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 442 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

Lys Leu Thr Thr Met Asp Trp Thr Trp Arg Phe Leu Phe Phe Val Val  
1 5 10 15

Ala Ala Ala Thr Gly Val Gln Ser Gln Val Gln Leu Val Gln Ser Gly  
20 25 30

Ala Glu Val Lys Lys Pro Gly Ser Ser Val Lys Val Ser Cys Lys Ala  
35 40 45

Ser Gly Gly Thr Phe Ser Ser Tyr Ala Ile Ser Trp Val Arg Gln Ala  
50 55 60

Pro Gly Gln Gly Leu Glu Trp Met Gly Gly Ile Ile Pro Ile Phe Gly  
65 70 75 80

Thr Ala Asn Tyr Ala Gln Lys Phe Gln Gly Arg Val Thr Ile Thr Ala  
85 90 95

Asp Glu Ser Thr Ala Arg Asp Asn Gly Ala Tyr Cys Ser Gly Gly Ser  
100 105 110

44

DRAFT - 08/02/2000

Cys Tyr Ser Gly Trp Phe Asp Pro Trp Gly Gln Gly Thr Leu Val Thr  
115 120 125

Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro  
130 135 140

Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val  
145 150 155 160

Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala  
165 170 175

Leu Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly  
180 185 190

Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Asp Lys  
195 200 205

Lys Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys  
210 215 220

Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro  
225 230 235 240

Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys  
245 250 255

Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Asn Phe Ser Trp  
260 265 270

Tyr Val Asp Gly Val Glu Val His Asn Asn Lys Thr Lys Pro Arg Glu  
275 280 285

Glu Asn Tyr Ser Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu  
290 295 300

His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Asn Val Ser Asn  
305 310 315 320

Lys Ala Leu Pro Ala Pro Ile Glu Lys Asn Ile Ser Lys Ala Lys Gly  
325 330 335

Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Asp Glu  
340 345 350

Leu Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr  
355 360 365

Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn  
370 375 380

Asn Tyr Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe  
385 390 395 400

Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn  
405 410 415

Val Phe Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr  
420 425 430

Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys  
435 440

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 42 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Pro Glu Met Leu Arg Asn Ser Thr Asp Thr Thr Pro Leu Thr Gly  
1 5 10 15

Pro Gly Thr Pro Glu Ser Thr Thr Val Glu Pro Ala Ala Arg Arg Ser  
20 25 30

Thr Gly Leu Asp Ala Gly Gly Ala Val Thr Glu  
35 40

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 16 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: not relevant
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Leu Thr Thr Glu Leu Ala Asn Met Gly Asn Leu Ser Thr Asp Ser Ala  
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 amino acids
- (B) TYPE: amino acids
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Thr Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser  
1               5                           10

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
- (B) TYPE: amino acids
- (C) STRANDEDNESS: not relevant
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Asp Tyr Glu Tyr Asp Glu Leu Pro  
1               5